Incorporating multiple genomic features with the utilization of interacting domain patterns to improve the prediction of protein-protein interactions

Abstract

Protein–protein interaction (PPI) networks play an outstanding role in the organization of life. Parallel to the growth of experimental techniques on determining PPIs, the emergence of computational methods has greatly accelerated the time needed for the identification of PPIs on a wide genomic scale. Although experimental approaches have limitations that can be complemented by the computational methods, the results from computational methods still suffer from high false positive rates which contribute to the lack of solid PPI information. Our study introduces the PPI-Filter; a computational framework aimed at improving PPI prediction results. It is a post-prediction process which involves filtration, using information based on three different genomic features; (i) gene ontology annotation (GOA), (ii) homologous interactions and (iii) protein families (PFAM) domain interactions. In the study, we incorporated a protein function prediction method, based on interacting domain patterns, the protein function predictor or PFP (), for the purpose of aiding the GOA. The goal is to improve the robustness of predicted PPI pairs by removing the false positive pairs and sustaining as much true positive pairs as possible, thus achieving a high confidence level of PPI datasets. The PPI-Filter has been proven to be applicable based on the satisfactory results obtained using signal-to-noise ratio (SNR) and strength measurements that were applied on different computational PPI prediction methods